

AUTOMATISIERUNG

Paperless-ngx: Dokumentenmanagement selbst hosten & automatisieren

Paperless-ngx als selbst gehostetes DMS für KMU: OCR, automatische Klassifikation, n8n-Integration und DSGVO-konforme Archivierung — ehrlicher Praxis-Guide für den DACH-Raum.

AUTOR

Natascha Reiner

VERÖFFENTLICHT

7. Mai 2026

ONLINE LESEN

<https://wissen.strukturaflow.it.com/paperless-ngx-dokumentenmanagement-selbst-hosten-automatisieren-der-kmu-gui/>

Der erste Schritt ist oft der schwierigste. Daten holen, Daten speichern — das klingt banal, ist es aber nicht. Sobald mehr als eine Person Dokumente bearbeitet, wird es schnell unübersichtlich und chaotisch. Und wenn man dann noch plant, digitale Unterstützung einzubauen — Automatisierung, KI, strukturierte Ablage — sitzt man plötzlich auf einem Fundament, das dafür gar nicht gebaut wurde. Dieser Artikel ist ein ehrlicher Einblick in Wissensmanagement für KMU: was funktioniert, was es wirklich kostet, und wie man es kosteneffizient aufsetzt, ohne sich in Tools zu verlieren.

Ein Handelsbetrieb in Klagenfurt — zehn Mitarbeitende, Lager, Warenein- und -ausgang, laufende Lieferantenkorrespondenz — kämpft täglich mit dem gleichen Problem: Lieferscheine landen als Scan im E-Mail-Postfach, Rechnungen kommen als PDF, Garantieunterlagen liegen als Papier im Ordner, und die Kollegin im Büro hat eine eigene Ordnerstruktur auf dem Desktop, die niemand sonst versteht. Wenn dann ein Gerät einen Garantiefall hat oder das Finanzamt eine Rechnung aus 2022 anfragt, beginnt die Suche. Ein einmal eingerichtetes Paperless-ngx auf dem eigenen Server hätte den Beleg in Sekunden per Volltextsuche gefunden und ihn beim Eingang bereits automatisch dem richtigen Lieferanten und Dokumenttyp zugeordnet — ohne dass die Kollegin irgendetwas manuell ablegen muss. Das klingt nach IT-Infrastruktur für Großunternehmen. Es ist tatsächlich Open Source, kostenlos und auf einem einfachen Linux-Server betreibbar.

Nach diesem Artikel wissen Sie, was Paperless-ngx kann, wo die Grenzen liegen, wie viel Einrichtungsaufwand realistisch ist — und wie Sie das System mit n8n so automatisieren, dass eingehende Rechnungen und Belege fast ohne manuellen Eingriff verarbeitet werden.

Was ist Paperless-ngx – und woher kommt es?

Paperless-ngx ist ein Open-Source-Dokumentenmanagementsystem (DMS), das physische und digitale Dokumente in ein durchsuchbares, strukturiertes digitales Archiv verwandelt. Die Kerntechnologie: automatische Texterkennung per OCR, Machine-Learning-gestützte Klassifikation und eine vollständige REST-API, über die sich das System mit anderen Tools verbinden lässt.

Die Geschichte des Projekts ist für Open Source typisch — und lehrreich. 2015 startete der kanadische Entwickler Daniel Quinn das ursprüngliche Paperless-Projekt. 2018 entstand der Fork Paperless-ng durch den deutschen Entwickler Jonas Winkler, der die Entwicklung erheblich vorantrieb. Als auch Winkler das Projekt nicht mehr alleine tragen konnte, formierte sich Anfang 2022 eine Community und übernahm das Projekt unter dem Namen Paperless-ngx. Das erste offizielle Release vereinte Beiträge von über 50 Entwicklern — Heise Online berichtete damals. Heute, 2026, befindet sich das Projekt stabil in der Versionsreihe 2.x und wird aktiv weiterentwickelt.

Warum ist das für KMU relevant? Weil Paperless-ngx heute kein Bastelprojekt mehr ist, sondern produktiv einsetzbares DMS — mit Mehrbenutzerverwaltung, Rollensystem, 2-Faktor-Authentifizierung und optionaler SSO-Integration über LDAP oder OpenID Connect.

Was Paperless-ngx wirklich kann – die Kernfunktionen im Überblick

OCR und automatische Texterkennung

Der Kern des Systems: Jedes eingehende Dokument wird automatisch durch Tesseract OCR verarbeitet. Das bedeutet, dass auch eingescannte PDFs — also Papierendokumente, die durch den Scanner als Bild vorliegen — vollständig durchsuchbar werden. Tesseract unterstützt über 100 Sprachen, Mehrsprachigkeit inklusive (relevant für DACH-Unternehmen mit Dokumenten in Deutsch, Englisch und weiteren Sprachen).

Ergebnis: Eine Rechnung von vor drei Jahren, die nur als Scan vorliegt, lässt sich über den enthaltenen Betrag, den Absender oder einen beliebigen Textfragment in Sekunden finden.

Automatische Klassifikation per Machine Learning

Paperless-ngx lernt aus Ihren Korrekturen. Das System schlägt bei neuen Dokumenten automatisch Korrespondenten (z.B. „Finanzamt Graz“), Dokumenttypen (z.B. „Rechnung“) und Tags vor — basierend darauf, wie Sie ähnliche Dokumente zuvor klassifiziert haben. Die Trefferquote steigt mit der Zeit und erreicht nach einem gewissen Datenbestand eine Qualität, bei der manuelle Nacharbeit zur Ausnahme wird. Wohin sich diese Technologie insgesamt entwickelt — von der reinen Texterkennung hin zu Systemen, die selbst über Folgeprozesse entscheiden — beleuchtet unser Ausblick Dokumenten-KI 2026: Vom Scanner zum autonomen Agenten.

Die Ablagestruktur ergibt sich automatisch aus konfigurierbaren Regeln. Statt manuell Ordner zu pflegen, definieren Sie Platzhalter — und Paperless-ngx generiert Pfade wie `2026/Finanzamt-Graz/Steuerbescheid/Einkommensteuerbescheid-2025.pdf` selbstständig.

Vier Dimensionen der Dokumentenorganisation

Paperless-ngx organisiert Dokumente entlang von vier Achsen:

DIMENSION	BEISPIEL	BESONDERHEIT
<u>Korrespondent</u>	Sparkasse Wien, Finanzamt Murau	Automatisch erkannt nach wenigen Beispielen
<u>Dokumenttyp</u>	Rechnung, Vertrag, Lohnabrechnung	Frei definierbar
Tags	Steuer, 2026, offen, Buchhaltung	Mehrere Tags pro Dokument möglich
Datum	Belegdatum, nicht Upload-Datum	Wird automatisch aus dem Inhalt extrahiert

Tags sind dabei besonders mächtig: Ein Dokument kann mehrere Tags tragen und so aus verschiedenen Blickwinkeln gefunden werden. Ein Inbox-Tag für neu eingegangene, noch nicht geprüfte Dokumente ist standardmäßig vorhanden.

Dokumenteneingang über mehrere Kanäle

Wie kommen Dokumente ins System? Fünf Wege funktionieren in der Praxis:

Consume-Ordner: Ein überwachtes Verzeichnis auf dem Server. Dateien, die dort abgelegt werden, verarbeitet Paperless-ngx automatisch. Scanner mit Netzwerk-Scan-Funktion (Scan-to-SMB/FTP) können direkt in diesen Ordner schreiben — der klassische Workflow für Büros mit Multifunktionsdruckern.

E-Mail (IMAP): Paperless-ngx kann Postfächer aktiv abfragen, Anhänge extrahieren und direkt verarbeiten. Filter nach Absender, Betreff oder Ordner sind konfigurierbar. Beispiel: Alle Mails von `rechnung@lieferant.at` → automatisch als Rechnung klassifiziert und mit Tag „Buchhaltung“ versehen.

Web-Upload: Direkter Drag-and-drop über die Browser-Oberfläche.

REST-API: Für die programmatische Anbindung — der Schlüssel zur Automatisierung über n8n und andere Tools (mehr dazu weiter unten).

Mobile: Die Web-Oberfläche ist responsive und mobil nutzbar; ein Foto eines Papierbelegs lässt sich direkt hochladen.

Übersichtstabelle der fünf Eingangswege für Dokumente in Paperless-ngx: von klassischen Scannern über E-Mail und Web-Upload bis zur API- und mobilen Integration.

Warum selbst hosten? DSGVO, GoBD und Datensouveränität

Das ist der entscheidende Unterschied zu cloudbasierten DMS-Lösungen wie DocuWare oder SharePoint Online: Bei Paperless-ngx auf dem eigenen Server verlassen Ihre Dokumente nie Ihren Einflussbereich.

Das ist keine ideologische Frage — es ist eine rechtliche und praktische. DSGVO und GoBD (Grundsätze zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen) stellen konkrete Anforderungen an die Aufbewahrung von Geschäftsdokumenten. Ein cloudbasiertes System, das Daten auf US-amerikanischen Servern speichert, erfüllt diese Anforderungen ohne zusätzliche Schutzmaßnahmen nach aktueller Rechtslage nicht.

Paperless-ngx unterstützt revisionssichere Archivierung: Dokumente werden intern im PDF/A-Format (dem Standard für Langzeitarchivierung) gespeichert und unveränderlich neben den Originaldateien abgelegt. Ein Audit-Log protokolliert alle Benutzeraktionen. Das reicht in vielen Fällen für GoBD-konforme Archivierung — eine Einschätzung durch einen Steuerberater oder Datenschutzbeauftragten für den konkreten Anwendungsfall bleibt trotzdem sinnvoll.

Ehrlicher Hinweis: Paperless-ngx ist kein zertifiziertes GoBD-System und wirbt auch nicht damit. Wer eine formal zertifizierte, revisionssichere Lösung mit Herstellergarantie braucht, muss tiefer in die Tasche greifen. Für die Mehrzahl der KMU, die ihre Ablage digitalisieren und strukturieren wollen, ist Paperless-ngx mehr als ausreichend — aber das sollten Sie mit Ihrem Steuerberater abstimmen.

Technisches Setup – was braucht man wirklich?

Paperless-ngx wird am besten als Docker-Container betrieben. Das Setup besteht aus drei Komponenten: dem eigentlichen Webserver, einer PostgreSQL-Datenbank und Redis als Message-Broker.

Systemanforderungen:

KOMPONENTE	MINDESTANFORDERUNG	EMPFOHLEN FÜR KMU-BETRIEB
RAM	2 GB	4–8 GB (mit OCR-Last)
CPU	2 Kerne	4 Kerne
Storage	Abhängig vom Dokumentenvolumen	mind. 50 GB, besser SSD
Betriebssystem	Jedes mit Docker-Support	Debian 12, Ubuntu 24.04

Hosting-Optionen:

Eigener Server (on-premise): Für KMU, die bereits einen Linux-Server betreiben, ist das der naheliegendste Weg. Volle Datenkontrolle, kein laufendes Abo, einmaliger Einrichtungsaufwand.

Strukturaflow Hosting — direkt aus der Steiermark: Wer sich nicht mit Server-Setup, Konfiguration und laufender Wartung befassen möchte, ist bei Strukturaflow richtig. Wir hosten Paperless-ngx für Unternehmen auf eigener, abgesicherter Infrastruktur — ohne dass Sie sich um Updates, Backups oder Verfügbarkeit kümmern müssen. Alles läuft in der Steiermark, unter österreichischem Recht, mit direktem Ansprechpartner. Kein Ticket-System, kein anonymer Support. Jetzt anfragen →

Paperless-ngx mit n8n automatisieren – die nächste Stufe

Das ist der Teil, der Paperless-ngx von einem reinen Archiv zu einem aktiven Bestandteil eines Automatisierungs-Workflows macht.

Paperless-ngx stellt eine vollständige REST-API bereit. Jede Aktion, die man über die Web-Oberfläche ausführen kann, lässt sich auch über die API auslösen: Dokument hochladen, Tags setzen, Korrespondent ändern, Dokument abrufen.

Verbindungsmöglichkeiten mit n8n

Community-Nodes: Es existieren zwei n8n Community-Nodes für Paperless-ngx, die Kern-Operationen (Dokumente lesen, Tags verwalten, Korrespondenten anlegen) ohne manuelle HTTP-Konfiguration ermöglichen.

HTTP-Request-Node: Für alle anderen Operationen: n8n's generischer HTTP-Node mit dem Paperless-API-Token in der Authorization-Header-Authentifizierung.

Webhooks (Paperless → n8n): Paperless-ngx unterstützt ausgehende Webhooks, die bei Ereignissen ausgelöst werden (z.B. „neues Dokument wurde verarbeitet“, „Tag X wurde gesetzt“). Das ermöglicht ereignisgesteuerte Automatisierungen.

Drei konkrete Automatisierungsszenarien

Szenario 1: Eingangsrechnungen aus E-Mail automatisch archivieren

Der Workflow: n8n überwacht ein IMAP-Postfach auf Mails mit PDF-Anhang von bekannten Lieferanten. Bei einem Treffer extrahiert n8n den Anhang und lädt ihn per API in Paperless-ngx hoch — inklusive vorgeordnetem Tag „buchhaltung“ und Korrespondent aus dem Absender der Mail.

- Einrichtungszeit: ca. 3–4 Stunden
- Voraussetzung: n8n-Instanz, Paperless-ngx mit zugänglicher API
- Einsparung: Manuelles Weiterleiten und Einpflegen von Rechnungen entfällt

Szenario 2: KI-gestützte Dokumententitel mit GPT

Ein offizielles n8n-Workflow-Template (verfügbar auf n8n.io) macht genau das: Wenn ein neues Dokument in Paperless-ngx eingeht und einen bestimmten Tag trägt (z.B. „ai-titling“), holt n8n den OCR-Text per API, bereinigt ihn von personenbezogenen Daten, schickt ihn an ein KI-Modell und schreibt den generierten Titel zurück in Paperless. Das Ergebnis: strukturierte Titel wie `[2026-05-07] - Rechnung - Stadtwerke Graz GmbH statt Scan_001.pdf`. Für weitere KI-Anwendungen im Büroalltag sehen Sie auch unsere [ChatGPT Prompts für Angebote schreiben](#).

- Einrichtungszeit: ca. 2 Stunden (Template vorhanden)
- Kosten: API-Kosten für GPT (bei typischem KMU-Volumen unter 5 €/Monat)
- Datenschutzhinweis: PII-Bereinigung vor dem KI-API-Aufruf ist im Template bereits vorgesehen — dennoch: Prüfen Sie, ob Ihre Dokumente besonders sensible Daten enthalten, bevor Sie externe KI-APIs einbinden. Mehr zur sicheren Nutzung im Artikel [DSGVO & ChatGPT im Unternehmen](#). Für sensible Branchen (Kanzleien, Gesundheitswesen) empfiehlt sich ein lokal betriebenes Modell über Ollama.

Szenario 3: Ablauf-Erinnerungen für Vertragsdokumente

Paperless-ngx unterstützt benutzerdefinierte Felder. Mit zwei Datumsfeldern (Ablaufdatum, Erinnerungsdatum) und einem täglichen n8n-Cron-Job lässt sich ein System bauen, das Sie automatisch per E-Mail benachrichtigt, wenn ein Vertrag bald ausläuft. Das funktioniert für Mietverträge, Wartungsverträge, Lizenzvereinbarungen — alles, was ein Enddatum hat.

- Einrichtungszeit: ca. 3–5 Stunden (inkl. benutzerdefinierte Felder in Paperless konfigurieren)
- Praxiswert: Hoch — vergessene Verlängerungsfristen kosten oft mehr als die Einrichtungszeit

Verbindung direkt auf dem Server

Für alle, die Paperless-ngx und n8n auf demselben Server betreiben: Beide Services kommunizieren über das Docker-interne Netzwerk — kein externer Port nötig, keine zusätzliche Absicherung erforderlich. Das ist der ideale Setup für datenschutzsensitive Umgebungen.

Für wen lohnt sich Paperless-ngx – und für wen nicht?

Gut geeignet für:

- Selbstständige und Freelancer mit wachsendem Dokumentenaufkommen (Rechnungen, Verträge, Korrespondenz)
- Kleine Betriebe (bis ~50 Mitarbeitende), die eine DSGVO-konforme Alternative zu Cloud-Ablage suchen
- Betriebe, die ohnehin einen Linux-Server betreiben und einen weiteren Service darauf deployen können
- Alle, die mit [n8n bereits arbeiten](#) und eine DMS-Komponente suchen

Weniger geeignet für:

- Betriebe ohne jegliche technische Ressource intern oder extern — das Setup ist kein One-Click-Installer
- Anwendungsfälle, die formal zertifiziertes, GoBD-geprüftes DMS mit Herstellergarantie erfordern (dann sind DATEV DMS, ELO oder DocuWare die richtigen Werkzeuge — mit entsprechendem Preisschild)
- Betriebe, die ein DMS mit nativer Integration in spezifische ERP-Systeme (z.B. SAP) benötigen

Realitätscheck Einrichtungsaufwand:

SZENARIO	EINRICHTUNGSZEIT	VORAUSSETZUNG
Paperless-ngx auf bestehendem Linux-Server	2–4 Stunden	Docker-Kenntnisse
Paperless-ngx auf neuem <u>Hetzner</u> VPS	4–6 Stunden	Linux-Grundkenntnisse
Erster n8n-Workflow (E-Mail → Paperless)	3–4 Stunden	Laufende n8n-Instanz
KI-Titelvergabe mit n8n-Template	2 Stunden	API-Key (OpenAI o. lokal)
Vollautomatisierter Rechnungseingang	6–10 Stunden	Alle obigen

DSGVO-Check: Was bei der Einrichtung zu beachten ist

Paperless-ngx selbst ist datenschutzneutral — die Verantwortung liegt beim Betreiber.

Die wichtigsten Punkte:

Dokumente bleiben auf dem eigenen Server — das ist per se DSGVO-freundlich. Problematisch wird es erst, wenn externe Dienste eingebunden werden: OCR-APIs von Drittanbietern, KI-APIs, Backup-Dienste auf US-Servern.

Empfehlungen für DACH-KMU:

- Hosting auf eigenem Server oder Hetzner/[IONOS](#)/[netcup](#) (alle deutsche Rechenzentren)
- Backups ebenfalls auf europäischen Servern oder lokal
- Vor dem Einbinden externer KI-APIs für [Dokumentenverarbeitung](#): prüfen, ob Dokumente personenbezogene Daten enthalten (fast immer der Fall bei Geschäftsdokumenten) — ggf. lokal laufendes Modell (Ollama auf dem eigenen Server) verwenden. Mehr zu [KI-Workflows mit Unternehmensdaten absichern](#) finden Sie in unserem dedizierten Artikel.
- [Multi-Faktor-Authentifizierung](#) aktivieren (in Paperless-ngx 2.x vorhanden)
- Reverse-Proxy mit TLS vor die Paperless-Instanz schalten, wenn sie über das Internet erreichbar ist (Nginx, Caddy oder Traefik)

Dieser Artikel ist keine Rechtsberatung. Für verbindliche Einschätzungen zur Datenschutz-Compliance empfiehlt sich eine Beratung durch einen Datenschutzbeauftragten. Mehr zum Thema [KI DSGVO-konform einsetzen](#) finden Sie in unserem ausführlichen KMU-Leitfaden.

Häufige Fragen zu Paperless-ngx

Ist [Paperless-ngx](#) eine Alternative zu SharePoint?

Für die Funktion Dokumentenarchivierung ja — für SharePoint als Kollaborationsplattform (gemeinsame Bearbeitung von Office-Dokumenten, Intranets) nein. Paperless-ngx ist ein spezialisiertes Archiv-DMS, kein allgemeines Dateiablagensystem.

Kann ich Dokumente aus [Nextcloud](#) in Paperless-ngx einpflegen?

Ja. Über die External-Files-Funktion lässt sich Nextcloud als Quelle anbinden. Der empfohlene Workflow: Nextcloud für aktive, täglich genutzte Dateien — Paperless-ngx als strukturiertes Langzeitarchiv für abgeschlossene Dokumente.

Wie viele Dokumente verwaltet Paperless-ngx ohne Performance-Probleme?

Mit ausreichend RAM (4–8 GB) und SSD-Storage verarbeitet Paperless-ngx auch fünfstelligen Dokumentenbestände stabil. Mehrere Nutzer berichten produktiven Betrieb mit über 50.000 Dokumenten auf einem einfachen VPS.

Was kostet die Einrichtung, wenn ich das nicht selbst mache?

Das hängt stark vom [Anbieter](#) ab. Managed-Hosting-Dienste starten ab etwa 10–15 €/Monat für kleine Installationen. Ein externer Dienstleister für Setup und initiale Konfiguration (inkl. n8n-Workflow) sollte mit 3–6 Stunden Aufwand kalkulieren.

Gibt es eine mobile App?

Eine offizielle native App gibt es nicht. Die Web-Oberfläche ist responsiv und auf dem Smartphone gut nutzbar. Für iOS und Android existieren Community-Clients (z.B. Paperless Mobile), die über die REST-API kommunizieren.

Nächste Schritte – wie Sie konkret anfangen

Ein pragmatischer Einstieg für KMU ohne bestehende DMS-Infrastruktur:

Schritt 1: Bestimmen Sie Ihren Dokumenteneingang. Wo entstehen die meisten zu archivierenden Dokumente — per E-Mail (Rechnungen als PDF), als Scan, oder beides? Das bestimmt, welche Eingangskanäle Sie zuerst konfigurieren.

Schritt 2: Entscheiden Sie über das Hosting. Haben Sie bereits einen Linux-Server in Betrieb? Dann Paperless-ngx dort als Docker-Container deployen. Keine eigene Hardware? Hetzner CX22, Docker-Setup, fertig.

Schritt 3: Starten Sie mit dem manuellen Upload, bevor Sie Automatisierung aufsetzen. Verstehen Sie, wie Tags, Korrespondenten und Dokumenttypen in Ihrer konkreten Ablagelogik funktionieren — das macht spätere Automatisierungsregeln deutlich besser.

Schritt 4: Erst wenn die manuelle Ablage läuft, den ersten n8n-Workflow aufsetzen. Der E-Mail-zu-Paperless-Workflow ist der beste Einstieg: hohe Frequenz, klare Logik, sofort messbarer Nutzen.

Paperless-ngx ist kein Selbstläufer — aber wer 4–6 Stunden investiert, hat eine Dokumenteninfrastruktur, die sich über Jahre ohne laufende Lizenzkosten amortisiert und vollständig unter eigener Kontrolle bleibt.

Sie haben keine Erfahrung mit Hosting und Server-Setup? Kein Problem.

Das Strukturaflow-Team ist spezialisiert auf genau das: IT-Infrastruktur für KMU, die funktioniert — ohne unnötige Komplexität und ohne aufgeblähte Lizenzkosten. Wir verkaufen keine Marken. Wir bauen Systeme, die zu Ihrem Betrieb passen.

Unser Fokus liegt auf Datensicherheit und effizientem Kosteneinsatz. Ob Paperless-ngx auf Ihrem Server, ein angebundener n8n-Automatisierungs-Workflow oder eine vollständige Dokumenteninfrastruktur von Grund auf — wir begleiten Setup, Konfiguration und laufenden Betrieb.

Jetzt unverbindlichen Termin buchen → strukturaflow.com/kontakt

NÄCHSTER SCHRITT

Mehr praktische KI-Anleitungen für KMU

Dieser Artikel ist Teil des KI-Hubs von Strukturaflow — einer deutschsprachigen Plattform für den praktischen KI-Einsatz in kleinen und mittleren Unternehmen.

<https://wissen.strukturaflow.it.com>